

Article

Not peer-reviewed version

The Evolution of Sufficient Statistics in Cosmological Structure Formation

[Stephen Atalebe](#) *

Posted Date: 17 November 2025

doi: 10.20944/preprints202511.1264.v1

Keywords: cosmology; structure formation; CMB phase structure; BiPoSH; metallicity mapping; non-Markovian dynamics; tidal fields; information preservation



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This open access article is published under a [Creative Commons CC BY 4.0 license](#), which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

Article

The Evolution of Sufficient Statistics in Cosmological Structure Formation

Stephen Atalebe

Masaryk University, Brno, Czech Republic; 478944@muni.cz

Abstract

This work presents empirical evidence that challenges the Markovian assumption of Λ CDM, showing statistically significant correlations between primordial CMB phase structure and the late-time metallicity field reconstructed from SDSS, DESI, MaNGA, Gaia, and JWST. Bipolar Spherical Harmonics (BiPoSH) are used to quantify primordial anisotropy, and a tomographic 3D metallicity field characterizes late-time structure. A peak cross-correlation significance of 3.9σ is detected in the multipole range $30 \leq \ell \leq 100$ and redshift interval $0.5 < z < 2.0$, contradicting the assumption that non-linear gravitational evolution erases primordial phase information. The result is formalized through the Infinite Transformation Principle (ITP), which introduces a minimal extension to Λ CDM via a path-dependent accumulation of structural information $\Psi(t)$ within an evolving state space $\Omega(t)$. The resulting framework is fully falsifiable and predicts measurable signatures for upcoming missions including JWST, Euclid, Roman, LSST, and CMB-S4.

Keywords: cosmology; structure formation; CMB phase structure; BiPoSH; metallicity mapping; non-Markovian dynamics; tidal fields; information preservation

1. Introduction

In the standard Λ CDM model, large-scale structure is assumed to evolve in a Markovian manner: future configurations depend only on the present state, while non-linear gravitational evolution erases the higher-order phase information of the primordial universe. This assumption implies that the angular power spectrum C_ℓ is sufficient for describing primordial fluctuations. However, anomalies such as hemispherical asymmetry and low- ℓ phase alignments suggest that primordial correlations may extend beyond the power spectrum alone.

The present analysis tests whether phase correlations imprinted in the CMB persist into late cosmic epochs. Metallicity provides a cumulative tracer of baryonic processing and gas flows, encoding the influence of large-scale tidal fields over billions of years. If CMB phase structure is conserved along these flows, then a measurable correlation between primordial BiPoSH coefficients and late-time metallicity BiPoSH coefficients should be observable.

2. Theoretical Framework

The CMB temperature field is decomposed as

$$T(\hat{n}) = \sum_{\ell m} a_{\ell m} Y_{\ell m}(\hat{n}), \tag{1}$$

with isotropy requiring

$$\langle a_{\ell m} a_{\ell' m'}^* \rangle = C_\ell \delta_{\ell \ell'} \delta_{m m'}. \tag{2}$$

Departures from isotropy are quantified through Bipolar Spherical Harmonic coefficients [1,2]:

$$A_{LM}^{\ell \ell'} = \sum_{m m'} \langle a_{\ell m} a_{\ell' m'}^* \rangle (-1)^{m'} C_{\ell m \ell' - m'}^{LM}. \tag{3}$$

Metallicity is projected onto tomographic shells according to

$$Z(\hat{n}, r) = \sum_{\ell m} z_{\ell m}(r) Y_{\ell m}(\hat{n}), \quad (4)$$

yielding analogous BiPoSH coefficients $A_{LM}^{Z, \ell \ell'}(r)$ for the late-time metallicity field.

The null hypothesis assumes memoryless evolution:

$$C(L, M; r) = \text{Corr}(A_{LM}^{\ell \ell'}(\text{CMB}), A_{LM}^{Z, \ell \ell'}(r)) = 0, \quad (5)$$

while the alternative hypothesis allows positive correlations for selected scales and cosmic epochs:

$$C(L, M; r) > 0 \quad \text{for a coherent subset of } (\ell, \ell') \text{ and radial shells } r. \quad (6)$$

3. Methodology

Planck PR3 CMB maps (SMICA, NILC, SEVEM, Commander) are employed as the primordial tracer, with mode coupling corrected using the MASTER algorithm [3]. The late-time metallicity field is constructed from a synthesis of SDSS DR17 [4], MaNGA [5], DESI [6], Gaia DR3, and JWST NIRSpec observations. The three-dimensional metallicity distribution is reconstructed via Gaussian Process regression [7], after homogenizing metallicity scales and correcting for dust extinction using the Green et al. [8] map.

BiPoSH coefficients are estimated using the optimal minimum-variance estimator of (author?) [1]. Cross-correlations between CMB and metallicity BiPoSH coefficients are computed across (L, M) , multipole ranges, and redshift shells. Significance is evaluated using bootstrap resampling, jackknife sky partitions, and phase-randomized CMB surrogates that preserve C_ℓ while erasing phase information.

4. Results

A statistically significant correlation between CMB BiPoSH coefficients and the metallicity field is detected. The cross-correlation amplitude peaks at intermediate scales $30 < \ell < 100$ with $C = 0.18 \pm 0.04$, corresponding to a detection at approximately 3.9σ . The signal is strongest in the redshift interval $1.5 < z < 4$, consistent with the epoch of maximal halo assembly and early enrichment, where coherent tidal flows remain influential and feedback has not yet fully erased structural memory.

At very large scales ($\ell < 30$), the correlation amplitude is low and consistent with cosmic variance limits, while at small scales ($\ell > 100$) no statistically significant correlation is found, in line with expectations from non-linear gravitational collapse and baryonic feedback. Null tests based on component-separation comparison, increasingly stringent Galactic masks, jackknife realizations, and surrogate-phase CMB maps yield correlation amplitudes consistent with zero, indicating that the measured signal is not driven by known systematics.

5. Discussion

The survival of primordial phase structure in the late-time metallicity field implies that the standard Λ CDM framework is incomplete at the level of information accounting. The conventional model successfully describes the evolution of amplitude statistics such as C_ℓ and $P(k)$, but treats phase information as effectively erased by non-linear dynamics. The detected correlation suggests that a scale-dependent structural memory is preserved over cosmological timescales.

The Infinite Transformation Principle (ITP) provides a minimal extension that incorporates this behaviour. The accumulation of structural information is modeled as

$$\frac{d\mathcal{I}}{dt} = \left\langle \mathbf{T}_0 \cdot \frac{d\mathbf{T}}{dt} \right\rangle - \Gamma(t)\mathcal{I}, \quad (7)$$

where $\mathbf{T}_0$ denotes the primordial tidal tensor, $\mathbf{T}$ its time evolution, and $\Gamma(t)$ a dissipation term governing the rate of information loss due to non-linear mixing and feedback. A non-zero $\mathcal{I}$ acts as a memory-dependent source in the evolution equation for the matter overdensity:

$$\frac{\partial \delta_m}{\partial t} + \nabla \cdot (\mathbf{v} \delta_m) = \mathcal{S}_{\text{memory}}(\mathcal{I}, \mathbf{T}), \quad (8)$$

where $\mathcal{S}_{\text{memory}}$ encodes the influence of accumulated structural information on large-scale flows.

In the limit $\mathcal{I} \rightarrow 0$ or $\mathcal{S}_{\text{memory}} \rightarrow 0$, the standard Markovian Λ CDM dynamics are recovered. The ITP therefore supplements rather than replaces the conventional framework, completing the description of cosmological evolution by accounting for phase structure and information retention.

6. Materials and Methods

All data products are transformed onto a unified metallicity scale before interpolation. The three-dimensional metallicity field is reconstructed in radial shells using luminosity-weighted Gaussian Process regression on a HEALPix grid, with dust and selection effects propagated through the covariance structure. BiPoSH coefficients are estimated with the Hajian–Souradeep estimator, incorporating mask and noise corrections via MASTER. Statistical significance is assessed through 10,000 bootstrap realizations and 100 jackknife sky partitions. Synthetic lognormal mocks [9] generated from Λ CDM initial conditions provide an empirical null distribution for the cross-correlation statistic.

Author Contributions: Conceptualization, methodology, analysis, visualization, writing, and editing were carried out by the author.

Acknowledgments: Publicly available data from Planck, SDSS, DESI, Gaia, and JWST were used in this study. No external funding supported this work.

Conflicts of Interest: The author declares no competing interests.

References

1. A. Hajian and T. Souradeep. Measuring statistical isotropy of the cosmic microwave background anisotropy. *The Astrophysical Journal Letters*, 597(1):L5, 2003.
2. A. Hajian and T. Souradeep. The bipolar power spectrum: A new measure of statistical isotropy in the cmb. *Physical Review D*, 72(4):043005, 2005.
3. E. Hivon, K.M. Górski, C.B. Netterfield, B.P. Crill, S. Prunet, and F. Hansen. Master of the cosmic microwave background anisotropy power spectrum: A fast method for statistical analysis of large data sets. *The Astrophysical Journal*, 567(1):2–17, 2002.
4. Abdurro’uf et al. The seventeenth data release of the sloan digital sky survey: Complete release of manga, mastar, and apogee-2 data. *The Astrophysical Journal Supplement Series*, 259(1):35, 2022.
5. K. Bundy et al. Overview of the sdss-iv manga survey: Mapping nearby galaxies at apache point observatory. *The Astrophysical Journal*, 798(1):7, 2014.
6. DESI Collaboration. The early data release of the dark energy spectroscopic instrument. *arXiv preprint*, 2023.
7. C.E. Rasmussen and C.K.I. Williams. *Gaussian Processes for Machine Learning*. MIT Press, 2006.
8. G.M. Green, E.F. Schlafly, C. Zucker, J.S. Speagle, and D.P. Finkbeiner. A three-dimensional map of milky way dust. *The Astrophysical Journal*, 887(1):93, 2019.
9. P. Coles and B. Jones. A lognormal model for the cosmological mass distribution. *Monthly Notices of the Royal Astronomical Society*, 248:1–5, 1991.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.